Catalogue no. 13F0026MIE — No. 001
ISSN: 1707-0783
ISBN: 978-0-662-46754-0

Research Paper

Pension and Wealth Research Paper Series

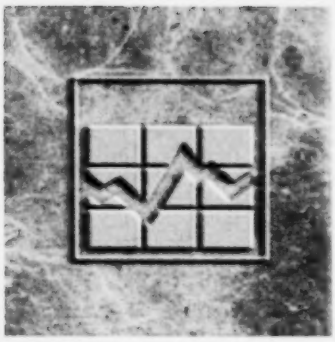
2005 Survey of Financial Security : Public Use Microdata User Guide

2005

by Pensions and Wealth Surveys Section

Income Statistics Division
Jean Talon Building, Ottawa, K1A 0T6

Telephone: 613-951-7355



Statistics Canada Statistique Canada

Canada

How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Income Statistics Division, Statistics Canada, Ottawa, Ontario, K1A 0T6 (telephone: 613-951-7355; 888-297-7355; income@statcan.ca).

For information about this product or the wide range of services and data available from Statistics Canada, visit our website at www.statcan.ca or contact us by e-mail at infostats@statcan.ca or by phone from 8:30am to 4:30pm Monday to Friday at:

Toll-free telephone (Canada and the United States):

Inquiries line	1-800-263-1136
National telecommunications device for the hearing impaired	1-800-363-7629
Fax line	1-877-287-4369
Depository Services Program inquiries line	1-800-635-7943
Depository Services Program fax line	1-800-565-7757

Statistics Canada national contact centre:	1-613-951-8116
Fax line	1-613-951-0581

Information to access the product

This product, catalogue no. 13F0026MIE, is available for free in electronic format. To obtain a single issue, visit our website at www.statcan.ca and select Publications.

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1-800-263-1136. The service standards are also published on www.statcan.ca under About us > Providing services to Canadians.

2008-03000/

Statistics Canada
Income Statistics Division

Pension and Wealth Research Paper Series

2005 Survey of Financial Security : Public Use Microdata User Guide

2005

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2007

All rights reserved. The content of this electronic publication may be reproduced, in whole or in part, and by any means, without further permission from Statistics Canada, subject to the following conditions: that it be done solely for the purposes of private study, research, criticism, review or newspaper summary, and/or for non-commercial purposes; and that Statistics Canada be fully acknowledged as follows: Source (or "Adapted from", if appropriate): Statistics Canada, year of publication, name of product, catalogue number, volume and issue numbers, reference period and page(s). Otherwise, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form, by any means—electronic, mechanical or photocopy—or for any purposes without prior written permission of Licensing Services, Client Services Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

September 2007

Catalogue no. 13F0026MIE
ISSN 1707-0783
ISBN 978-0-662-46754-0

Frequency: occasional

Ottawa

La version française de cette publication est disponible sur demande (nº 13F0026MIF au catalogue).

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Table of contents

1.0 Introduction 5

1.1 Overview 5
1.2 History 5
1.3 How to cite SFS in publications 5

2.0 Key findings 7

3.0 Concepts and definitions 9

3.1 Net worth 10
3.2 Assets 10
3.3 Debts 12
3.4 Family type 13

4.0 Survey methodology 14

4.1 The survey universe 14
4.2 Survey content and reference period 14
4.3 The sample 15
4.4 Data collection 15
4.5 Data processing and quality control 15
4.6 Imputation of missing data 15
4.7 Weighting 16

5.0 Data accuracy and quality 18

5.1 Sampling errors 18
5.2 Standard error, confidence intervals and coefficient of variation 18
5.3 Non-sampling errors 21
5.3.1 Coverage errors 21
5.3.2 Response errors 21
5.3.3 Non-response errors 22
5.3.4 Processing errors 22
5.3.5 Estimation errors 22
5.3.6 Analysis error 22
5.4 Treatment of large values 22
5.5 Impact of sampling and non-sampling errors on SFS estimates 23
5.6 Comparability of data and related sources 23
5.7 Response rates 23

6.0 Record layout, data dictionary and univariate distributions 25

6.1 Columns of the record layout 25
6.2 Data dictionary 25
6.3 Univariate distribution 25
6.4 Reserved Codes 26

7.0 Guidelines for analysis and presentation 27

7.1 Applying weights 27
7.2 Rounding guidelines 27
7.3 Missing values and reserved codes 27

8.0 Guidelines for release 28

8.1 Introduction 28
8.2 Minimum sizes of estimates for release 28
8.3 Hypothesis tests provided by statistical software packages 28

9.0 Appendix A – SFS production totals VS. PUMF totals 30

10.0 Appendix B – SFS totals 31

1.0 Introduction

1.1 Overview

The 2005 Survey of Financial Security (SFS) provides a comprehensive picture of the net worth of Canadians. Information was collected on the value of all major financial and non-financial assets and on the money owing on mortgages, vehicles, credit cards, student loans and other debts. The value of these assets less the debts is referred to as net worth.

The cross-sectional public use microdata file (PUMF) is a collection of income, expenses, assets, debts and wealth data on Canadian families. This file contains information collected from more than 5,200 family units residing in private households in Canada. All records have been thoroughly screened to ensure the anonymity of respondents.

This manual was produced as a reference guide to help users manipulate the microdata file of the survey results.

For more information, or to enquire about concepts, methods or data quality, please contact:

Income Statistics Division
Toll-free 1-888-297-7355 or 613-951-7355
income@statcan.ca

1.2 History

Since the 1950s, Statistics Canada has conducted occasional surveys on the assets and debts of Canadians. Up to, and including, 1984, these surveys were supplements to the more regular income surveys, known as the Surveys of Consumer Finances (SCF). Originally, these surveys covered non-farm households and a limited list of assets and debts. The coverage has since expanded to include almost all private households in Canada and nearly all marketable and pension assets.

Over the six-year period between 1999 and 2005, a number of important factors influenced the evolution of the wealth distribution in Canada. The real estate market experienced strong growth over the period, with historically low interest rates and favorable economic conditions spurring new construction and inflating the value of existing homes.

The Canadian stock market plunged in 2002 in the aftermath of the high-tech bubble burst, temporarily wiping out billions in family net worth. It had fully recovered by 2005 on the strength of a booming resource sector.

With the cost of borrowing at all-time lows, consumer debt rose to unprecedented levels, while debt service ratios nonetheless have not increased in relation to income. The banking sector targeted an increasing portion of its lending activity at consumers, making consumer credit more available in a broader variety of forms.

1.3 How to cite SFS in publications

For publication of any information based on the SFS microdata file on CD-ROM (13M0006XCB), the following form of accreditation is recommended:

"This analysis is based on Statistics Canada's Survey of Financial Security Public Use Microdata, 2005, which contains anonymous data collected in the Survey of Financial Security. All computations on these

microdata were prepared by (Name of user). The responsibility for the use and interpretation of these data is entirely that of the author(s)".

2.0 Key findings

The 2005 survey results showed that:

- The median net worth of Canada's estimated 13.3 million family units amounted to $148,400 in 2005, a 23.2% increase from 1999, after adjusting for inflation.

- Gains in median net worth were seen in all but the lowest wealth quintile since 1999. Increases were highest in the fourth and top wealth quintiles. This signifies increased inequality in Canada's wealth distribution. The 20% of family units with the highest net worth held 69.2% of all personal wealth in the country in 2005.

- Total assets, everything from pension assets, stocks and bonds to principal residences, amounted to over $5.6 trillion, nearly 1.4 times the estimate of $3.9 trillion in 1999.

- The single most important asset for Canadians was their principal residence, accounting for one-third of the total. This was followed by employer pension plans (EPP), which represented almost 18.5% of all assets.

- The increase in the market value of real estate was the major contributor to the growth of total assets of Canadian families between 1999 and 2005. Real estate accounted for almost 50% of the increase with gains largely attributable to price increases as well as an increase in the number of families holding this type of asset (+13.6%).

- A significant change in the composition of assets between 1999 and 2005 was the growth in the amount invested in real estate such as cottages, timeshares, rental properties and other commercial properties. The aggregate amount in this type of real estate was roughly 1.8 times larger in 2005 than in 1999 amounting to almost $481 billion from $266 billion, in constant 2005 dollars. This was by far the largest rate of growth of any single asset.

- Assets held in private pension instruments such as employer pension plans, Registered Retirement Savings Plans (RRSPs) and Registered Retirement Income Funds (RRIFs), were the second largest contributor. The increase in this category was largely driven by the sharp growth in the value of EPPs.

- About 9.4 million family units, or 70.6% of the total, had some form of pension assets in 2005, whether they were EPPs, RRSPs or RRIFs. In total, pension assets peaked for family units in which the major income recipient was aged between 55 and 64.

- In 2005, these family units, which would have been approaching retirement or just recently retired, had total median pension assets of $242,500. This was well above the median for all family units of $68,000.

- About 58% of family units had RRSPs, RRIFs or Locked-in Retirement Accounts (LIRAs) in 2005. The median value of these plans was $30,000.

- Just under one-half (48.6%) of family units had assets in employer pension plans. The median value of those assets was a much larger $68,300.

- Canadians had debts estimated at $760 billion in 2005, three-quarters of which took the form of mortgages. This represented a 43.3% increase from 1999. This increase can be explained by both the rise in cost to purchase a home and the increase in the number of families who owned their principal residence with a mortgage (+16.6%). The median value of mortgages in 2005 amounted to $93,000, up 17.0% from about $79,500 in 1999.

- Debts from lines of credit amounted to roughly $68 billion in 2005, about 9% of the total. Loans on owned vehicles amounted to about $46 billion, or 6.1% of the total. Canadian family units held almost $25.8 billion in outstanding credit card and instalment debt, or 3.4% of the total, while student loans approached $20 billion.

- Outstanding credit card and instalment debt was up 58.4% from $16.3 billion in 1999. The median credit card and instalment debt rose to $2,400 in 2005.

- Almost 11 million Canadian families reported owning a credit card in 2005. Of the over 2 million family units that reported not owning one, nearly 19% of these families were refused this type of credit. Nearly 73% of families who had credit cards reported they pay off their balances each month. The median credit limit on all credit cards owned was $10,000.

- A notable development over the six-year period was the growth in outstanding line-of-credit debt, which surged 2.3 times. At the same time, the number of family units with line-of-credit debt increased almost 77% to 3.3 million. Almost a quarter (24.9%) of all families had a line-of-credit debt in 2005, compared with 15.4% in 1999. The median debt in this form rose from $5,800 to $9,000. Much of the increase was secured by residential assets in the form of home equity lines of credit.

- The median debt load for family units rose 37.8% from $32,300 to $44,500 between 1999 and 2005. Canadians had an estimated $13.52 in debts for every $100 in assets in 2005, up from $13.06 in 1999. However, the debt burden was much higher for certain family types, such as lone-parents families ($28.32 per $100), and couples with children under 18 ($20.03 per $100).

- Family units headed by someone under 35 carried the highest debt load, $39.40 for every $100 of assets, up 17.2% when compared to 1999. Family units headed by seniors carried the least amount of debt, $2.26 per $100 of assets.

- The survey found that of the 7.4 million people aged 55 and over, 63.7% reported having been retired.

- Respondents were asked to report all their reasons for retiring. The top three Canadians cited were: personal or family responsibilities (23.7%), health (22.8%), and having completed the required number of years to qualify for a pension (19.6%).

- Of the roughly 4.7 million people who retired, just over 784,000 of them returned to work for pay following their first retirement. Over 48% of those returning to work for pay reported that financial considerations were the most common reason for returning to work. Other significant reasons cited were that work was interesting and job offers from employer.

3.0 Concepts and definitions

This chapter outlines the definitions of the main assets, debts and wealth concepts and their components.

Table 3-1 below illustrates the components of the net worth calculation accounted for by the Survey of Financial Security. The value of all assets less all debts is net worth. A family's net worth can be thought of as the amount of money they would have if they liquidated their assets and paid off all of their debts. The PUMF variable names appear in brackets.

Table 3-1 Net worth calculation

Assets **(WATOTPT WATOTPG)**	**less:**	**Total debts** **(WDTOTAL)**	**equals**	**Net worth** **(WNETWPT WNETWPG)**
RRSPs/LIRAs **(WARRSPL)**		Mortgage on principal residence **(WDPRMOR)**		
RRIF **(WARRIF)**		Mortgage on other real estate **(WDSTOMOR)**		
Employer Pension Plans (EPP) **(WARPPT, WARPPG)**		Line of credit **(WDSTLOC)**		
Retirement funds, other **(WAOTPEN)**		Credit card and instalment debt **(WDSTCRED)**		
Deposits in financial institutions **(WASTDEPT)**		Student loans **(WDSLOAN)**		
Mutual funds and other investment funds **(WASTMUIC)**		Vehicle loans **(WDSTVHLN)**		
Stocks **(WASTSTCK)**		Other debt **(WDSTODBT)**		
Bonds **(WASTBOND)**				
Other financial assets, non-pension **(WASTOINP)**				
Principal residence **(WAPRVAL)**				
Other real estate **(WASTREST)**				
Vehicles **(WASTVHLE)**				
Other non-financial assets **(WASTONOF)**				
Equity in business **(WBUSEQ)**				

3.1 Net worth

The net worth (sometimes referred to as wealth) of a family unit is defined as the difference between the value of its total asset holdings and the amount of total indebtedness .

There are two types of net worth variables:

1. **WNETWPT** – Net worth of the family unit. (Assets including current pensions valued on termination basis[1] (WATOTPT) - debts (WDTOTAL).
2. **WNETWPG** – Net worth of the family unit. (Assets including current pensions valued on going concern basis[1] (WATOTPG) - debts (WDTOTAL).

Respondents were asked to provide the value of the asset or the amount of the debt at a time as close as possible to the date of the interview. Assets and debts were reported for the family unit as a whole and not for each person in the family. The assets and debts included in the survey are identified below.

3.2 Assets

Respondents were asked to report the market value of the asset that is the amount they would receive if they had sold the asset at the time of the survey. If available, respondents were encouraged to consult financial records. When the value could not be determined through an independent source, the respondent was asked to estimate the value. This is in itself prone to error. In the case of vehicles, respondents were asked to provide the make, model and year in addition to the estimated value. This information was used to impute for non-response and also to identify potential reporting errors. Values provided by respondents were not adjusted unless they were judged to be an error, resulting, for example, from data entry. If the respondent either over or underestimated the value of an asset by a relatively small proportion, this would not be readily apparent. However, extreme values were reviewed and adjusted if necessary.

The value of all invested assets was to include accrued earnings or interest. Respondents were asked to estimate the actual value, at the time of the survey. In one case, for the value of the contents of the principal residence, the respondent was able to select one of 16 ranges.

The definitions of the assets items identified in table 3-1 are:

Assets, total (WATOTPT, WATOTPG): Total value of all financial assets, non-financial assets and equity in business.

There are two types of total asset variables:

1. **WATOTPT** - Total assets, including employer pension plans (current plans valued on termination[1] basis).
2. **WATOTPG** - Total assets, including employer pension plans (current plans valued on a going concern[1] basis)

Bonds (WASTBOND): Bonds are the total value, including earnings, of federal and provincial savings bonds and other bonds issued by governments and corporations. Includes investment in foreign bonds but excludes the amount held within registered plans.

Deposits (WASTDEPT): Deposits are the total amount, including interest, of all chequing and savings accounts with a non-zero balance and of other deposits such as term deposits and Guaranteed Investment Certificates. These amounts would generally be held in financial institutions such as chartered banks, trust companies, co-ops and caisses populaires. This item includes only the amount held outside of registered plans.

1. Employer pension plan valuation is explained further in this chapter.

Employer pension plans (WARPPT, WARPPG): An employer pension plan (EPP) is an employer-sponsored plan registered with Canada Customs and Revenue Agency and most commonly also with one of the pension regulatory authorities. The purpose of such plans is to provide employees with a regular income at retirement.

There are two commonly used approaches to valuing EPP assets: the **going concern** and the **termination** approach. The two EPP variables included on the PUMF are:

1. **WARPPT** - Current pension plans valued on a termination basis.
2. **WARPPG** – Current pension plans valued on a going concern basis

The main differences between the two valuation methods are:

(a) Although future service is not considered in either type of valuation, in a going concern valuation assumptions are made about future salary increases. As many EPPs base the amount of the pension on average earnings close to the time of retirement, assuming salary increases up to that time will obviously increase the value of the benefit. In a termination valuation, salary increases are not considered.

(b) Interest rates for a termination valuation are assumed based on current market rates. For a going concern valuation longer term interest rates are assumed.

(c) The going concern valuation method is applicable only for current members of certain types of EPPs. Those with deferred pensions (people who had previously belonged to an EPP) and those receiving benefits are no longer members of the plan so future salary increases need not be considered.

When analyzing SFS data the termination valuation approach is generally used. That approach is more consistent with the basis on which other assets are valued, in that future expectations are not taken into consideration and current market conditions are used to estimate the value. The termination approach, however, can underestimate the value of the benefit earned (accrued) as of the time of the survey because many employees will continue to participate in the plan, and therefore receive a pension based on their salary closer to the time of retirement. In order to allow users the option of selecting the value of the EPP that is most appropriate for their type of analysis both values have been produced and are available.

In valuing benefits for those respondents who belonged to a pension plan at the time of the survey, only plan membership up to the time of the survey has been considered. Therefore, in the case of a person who was 45 at the time of the survey and who had participated in an EPP for 10 years, the pension would be valued for the 10 years of known service.

For more information on employer pension valuation see M. Cohen, H. Frenken and K. Maser, *Survey of Financial Security: Methodology for estimating the value of employer pension plan benefits*, Statistics Canada, Catalogue 13F0026MIE-01003.

Equity in business (WBUSEQ): The estimated amount the respondent would receive if the business were sold, after deducting any outstanding debts to be paid.

Locked-in Retirement Account (included in **WARRSPL**): A Locked-In Retirement Account (LIRA) is an RRSP in which the money is locked-in until the person reaches a specified age. LIRAs are included in the RRSP category. This money would have been transferred from an employer pension plan after the individual terminated employment. For the most part, LIRAs came into use in the late 1980s, when revisions to pension regulatory legislation provided for enhanced portability of pension accruals on termination of employment.

Mutual funds and other investment funds (WASTMUIC): The total value, including investment earnings, of all holdings in mutual and investment funds. Excludes the amount held within registered plans.

Principal residence (WAPRVAL): Market value, as estimated by the respondent, of the residence where the respondent lives. If the respondent has two residences, this would be the one where they most often live. If the respondent shares ownership of the home with someone outside the family, only the family's share is included. If the property is a farm, the estimated value of the farmhouse is included; the value of the farmland would be included either with business equity or with other real estate, if no business were reported.

Real estate, other (WASTREST): Estimated market value of real estate other than the respondent's principal residence. Included would be second homes, vacation homes, timeshares, rental property (residential or non-residential) or vacant lots. Real estate includes property in Canada or outside.

Registered Retirement Savings Plans (included in **WARRSPL**): A Registered Retirement Savings Plan (RRSP) is a capital accumulation program designed to encourage saving for retirement. Contributions are tax-deductible within prescribed limits. Investment income earned in the RRSP is tax exempt, but benefits are taxable.

The RRSP could be held in deposits, mutual funds, stocks or bonds. As well, this includes the amount held in Locked-in retirement accounts (LIRAs); see definition above.

Registered retirement income funds (WARRIF): A Registered Retirement Income Fund (RRIF) is intended to provide a regular income in retirement. Monies in RRSPs must be transferred to a RRIF or an annuity before the end of the year in which the owner of the RRSP turns 71[2]. Payments from an RRIF may be varied, but a minimum amount must be withdrawn annually. Also includes monies in locked-in retirement income funds (LRIFs) and life income funds (LIFs); these plans are intended to receive amounts transferred from an employer pension plan.

Stocks (WASTSTCK): Total value, including earnings, of all publicly-traded common and preferred shares. Includes foreign stock but excludes the amount held within registered plans.

Vehicles (WASTVHLE): Estimated value of cars, trucks, vans, sport utility vehicles as well as motorcycles, mobile homes, boats and snowmobiles. Excludes vehicles owned by the respondent's business and vehicles that are leased.

3.3 Debts

The amount reported for debts is not intended to include interest owing, as this would most often not be known.

The debt items listed in table 3-1 comprise the following:

Debts, total (WDTOTAL): Total of all debts for the family unit.

Credit card and instalment debt (WDSTCRED): For credit cards, the amount owing on the last bill, excluding any new purchases. Includes major credit cards (VISA, Mastercard, American Express, Diners Club/en Route) and retail store cards, gasoline station cards, etc. Instalment debt is the total amount owing on deferred payment or instalment plans where the purchased item is to be paid for over a period of time.

Line of credit (WDSTLOC): Total amount owing on both a home equity line of credit and a regular line of credit. This does not refer to the credit limit on the LOC.

2. At the time the SFS was conducted, the age was 69.

Mortgages, on principal residence (WDPRMOR): Outstanding amount owing on the respondent's principal residence. If the respondent shares ownership of the home with someone outside the family, only the family's share of the mortgage is included. If the property is a farm, the mortgage owing on the farmhouse is included; the mortgage on the remainder of the farm would implicitly be included with business equity or would be included with mortgage owing on other real estate, if no business were reported.

Mortgages, on other real estate (WDSTOMOR): Respondent's share of the mortgage owing on second homes, vacation homes, timeshares, rental property (residential or non-residential) or vacant lots.

Student loans (WDSLOAN): Amount owing on loans taken out to attend a post secondary education program. These loans are most often taken through the Canada Student Loan Program or one of the provincial student loan programs. This item also includes amounts owing on loans taken directly from a financial institution to attend school.

Vehicle loans (WDSTVHLN): Amount owing on loans for those vehicles listed under assets.

3.4 Family type

Within the family type classification, the following definitions apply:

Couples: Couples include legally married, common-law and same-sex relationships.

Couples with children: Couples living with a child or children (by birth, adopted, step, or foster) under age 18. Children aged 18 or over are considered to be "other relatives". Other relatives may also be in the family.

Economic family: An economic family is defined as a group of two or more persons who live in the same dwelling and are related to each other by blood, marriage, common law or adoption.

Elderly/elderly families: Person aged 65 and over. In the case of elderly families, the major income recipient is aged 65 and over.

Family units: Includes economic families of two or more and unattached individuals.

Lone-parent families: One parent living with at least one child under age 18. Families where the parent is 65 years and older are excluded.

Other non-elderly families: Couples living with a child or children (biological, adopted, step, or foster) aged 18 or over and/or with other relatives, but not living with a child or children under the age of 18. Also includes lone parent families (with children of all ages) and related persons (e.g., siblings, cousins) living together.

Unattached individual: An unattached individual is a person living either alone or with others to whom he or she is unrelated, such as roommates or a lodger.

Major income recipient or earner: For each family, the major income recipient is the person with the highest income before tax. For persons with negative total income before tax, the absolute value of their income is used, to reflect the fact that negative incomes generally arise from losses "earned" in the market-place and are not meant to be sustained. In the rare situations where two persons have exactly the same income, the older person is the major income recipient.

4.0 Survey methodology

4.1 The survey universe

The 2005 Survey of Financial Security was carried out in all ten provinces, the territories were not included. Those living on Indian reserves and crown lands and official representatives of foreign countries living in Canada and their families were also excluded from the survey. Members of religious and other communal colonies, members of the Canadian Forces living in military camps and people living in residences for senior citizens were excluded, as were people living full time in institutions, for example, inmates of penal institutions and chronic care patients living in hospitals and nursing homes. The survey covers about 98% of the population in the ten provinces.

Information was not gathered from persons temporarily living away from their families (for example, students at university) because it would be gathered from their families if selected. In this way, double counting of such individuals was avoided.

4.2 Survey content and reference period

With a few exceptions, the reference period for the information was the time of data collection (May to July 2005). For the asset and debt information respondents were asked to provide an estimate of the value or amount as close to the survey date as possible, recognizing that their most recent statement may have been as of the end of the previous calendar year, or for the last financial quarter.

Some of the information was collected for each person in the family 15 years of age and over. The assets and debts, however, were collected for the family as a whole, because they often cannot easily be assigned to one person in the family. Specifically, the following information was collected:

From each family member 15 years of age and over:

- demographics (age, sex, marital status);
- ethno-cultural characteristics;
- education;
- current employment;
- income, for the calendar year 2004.

From each family member 25 years of age and over:

- previous employer pension plans
- pension plan benefits

From each family member 45 years of age and over:

- retirement information

For the family unit as a whole:

- financial and non-financial assets;
- equity in business;
- debt in the form of mortgages, vehicle loans, credit card and line of credit debt, student loans and other debt.
- distribution of registered plans investments
- distribution of mutual funds investments

4.3 The sample

The total sample for the 2005 Survey of Financial Security was 9,000 dwellings; it was drawn from two sources.

The main sample, drawn from an area frame, consisted of 7,500 dwellings. This area sample was a stratified, multi-stage sample selected from the Labour Force Survey (LFS) sampling frame. Dwellings selected for this survey had not previously participated in a labour force or financial survey conducted by Statistics Canada. Sample selection comprised three steps: the selection of clusters (small geographic areas) from the LFS frame, field listing of all addresses within each selected cluster, and the selection of dwellings within these selected clusters. At the time that the SFS sample was selected the LFS frame was using 2001 Census geography.

The second portion of the sample, 1,500 dwellings, was drawn from geographic areas in which a large proportion of family units had what was defined as "high-income". This sample was included to improve the representation in the sample of high income families, as a disproportionate share of net worth is held by such higher-income family units. For purposes of this sample the income cutoff was total family income of at least $200,000 or investment income of at least $50,000. The latter was used to take into account those family units that may not have high income from employment but have substantial assets that generate investment income.

4.4 Data collection

Data were collected during a personal interview using a paper questionnaire.

For families, the interview was held with the family member with most knowledge of the family's financial situation. If necessary, follow-up was done with other family members. Proxy response was accepted. This allowed one family member to answer questions on behalf of any or all other members of the family, provided he or she was willing and able to do so.

To reduce response burden, for the questions on 2004 income, respondents could give Statistics Canada permission to use the income information from their T1 tax return. Close to 80% of survey respondents gave their consent to use these administrative records.

4.5 Data processing and quality control

In-house scanning software was used to capture survey data from the questionnaire. A quality control operation was applied to ensure that pre-specified quality standards were achieved. Data then passed through an automated edit system to identify inconsistencies and potential errors in the data.

4.6 Imputation of missing data

Missing responses were imputed for all key fields in the questionnaire. Where possible, information was imputed deterministically, using other information reported by the respondent. For example, when the respondent could not estimate the value of their vehicle, the reported make, model and year was used to impute a value. This value was determined by consulting a reference book. When deterministic imputation was not possible, hotdeck imputation methods were used in most cases, and for all components of income and net worth, nearest neighbour techniques were employed. These methods involve identifying another individual or family with similar characteristics to become the "donor" and provide the imputed value. Income data obtained from tax returns are considered complete and thus do not require imputation.

Table 4-1 shows the percentage of asset and debt values that was determined through imputation.

Table 4-1 Imputation rates

	Assets or debts (after imputation)[1]	Imputed[1]
	%	
Assets	100	25
Pension assets	...	...
RRSPs / LIRAs / RRIFs / other	11	9
Employer pension plan (EPP)[2]	...	...
Financial assets, non-pension	10	14
Deposits in financial institutions	4	12
Mutual funds/investments funds/income trusts	2	12
Stocks	2	15
Bonds	1	13
Other financial assets	1	17
Non-financial assets	50	4
Principal residence	33	4
Other real estate	9	7
Vehicles	3	1
Other non-financial assets	5	3
Equity in business	11	16
Debts	100	3
Mortgages	75	3
Principal residence	64	3
Other real estate	11	2
Line of credit	9	4
Credit card and instalment debt	3	3
Student loans	3	3
Vehicle loans	6	5
Other debt	4	1
Net worth	...	28

... not applicable

1. This means, for example, that the principal residence (the home) constituted 33% of total assets and that 4% of the total amount for principal residence was imputed.
2. For employer pension plans, all values have been derived and not imputed.

4.7 Weighting

The estimation of population characteristics from a survey is based on the premise that each sampled unit represents, in addition to itself, a certain number of unsampled units in the population. A basic survey weight is attached to each sample record to indicate the number of units in the population that it represents. Adjustments are then applied to the basic survey weights in order to improve the reliability of the estimates.

The basic weights are first inflated to compensate for non-response. This adjustment was applied within groups of sample units that are geographically close and the area and high income samples were adjusted separately.

A frame allocation factor was then applied. Since the high income frame overlaps completely with the area frame, units on the high income frame were eligible to be selected from either the high income or the area frame. The frame allocation factor was applied to non-response adjusted weights to account for this increased probability of selection.

The weights are then further adjusted to ensure that estimates of relevant population characteristics would respect known population totals from sources external to the survey. The population totals used for the SFS were based on Statistics Canada's Demography Division population counts for different province - age - sex groups. The weights were also adjusted to ensure that the number of 1-person and 2-person families, and the number of 1-person and 2-person households agreed with known totals by region.

Additionally in 2005, two new sources of weight adjustments were introduced. The first adjustment was based on administrative data from the T4 file. Weight adjustments were made to ensure that the survey distribution of earnings reflected approximately the same distribution as the T4 population. The second new adjustment made use of Survey of Labour and Income Dynamics (SLID) data to improve estimation. SFS as the smaller sample survey borrowed strength from SLID, the larger sample survey to not only improve SFS estimates but also to bring estimates for the 2 surveys more in line with each other.

5.0 Data accuracy and quality

5.1 Sampling errors

Sampling errors arise from estimating a population characteristic by looking at only one portion of the population rather than the entire population. It refers to the difference between the estimate derived from a sample survey and the 'true' value that would result if a census of the whole population were taken under the same conditions. There are no sampling errors in a census because the calculations are based on the entire population.

5.2 Standard error, confidence intervals and coefficient of variation

A common measure of sampling error is the standard error (SE). The standard error measures the degree of variation introduced in estimates by selecting one particular sample rather than another of the same size and design. The standard error may also be used to calculate confidence intervals associated with an estimate (Y).

Confidence intervals (CI) are used to express the precision of the estimate. It has been demonstrated mathematically that, if the sampling were repeated many times, the true population value would lie within the Y +/- 2SE confidence interval 95 times out of 100 and within the narrower confidence interval defined by Y +/- SE, 68 times out of 100.

Another important measure of sampling error is given by the coefficient of variation (CV). The coefficient of variation is the standard error of an estimate, expressed as a ratio or percentage of the estimate (i.e. 100 x SE / Y).

To illustrate the relationship between the standard error, the confidence intervals and the coefficient of variation, let us take the following example. Suppose that the estimated median net worth from a given source is $10,000, and that its corresponding standard error is $200. The coefficient of variation is therefore equal to 2%. The 95% confidence interval estimated from this sample ranges from $9,600 to $10,400, i.e. $10,000 +/- $400. This means that with a 95% degree of confidence, it can be asserted that the median net worth of the target population is between $9,600 and $10,400.

Estimates with a coefficient of variation less than 16.6% are considered reliable for general use. Estimates with coefficients of variation between 16.6% and 33.3% should be accompanied by a warning to caution users about the high levels of error. Estimates with coefficients of variation higher than 33.3% are deemed to be unreliable. For estimates of net worth in this survey, CVs greater then 33.3% generally occur when the sample size contributing to an estimate is 25 or less. This affects the level of detail in published tables and, in particular, limits the availability of provincial statistics.

Table 5-1 provides quality level guidelines used at Statistics Canada.

Table 5-1 Quality level guidelines

Quality level of estimate	Guidelines
1) Acceptable	Estimates have a sample size of 30 or more, and low coefficients of variation in the range of **0.0% to 16.5%.** No warning is required.
2) Marginal	Estimates have a sample size of 30 or more, and high coefficients of variation in the range of **16.6% to 33.3%.** Estimates should be flagged with the letter M (or some similar identifier). They should be accompanied by a warning to caution subsequent users about the high levels of error, associated with the estimates.
3) Unacceptable	Estimates have a sample size of less than 30, or very high coefficients of variation in excess of **33.3%.** Statistics Canada recommends not to release estimates of unacceptable quality. However, if the user chooses to do so then estimates should be flagged with the letter U (or some similar identifier) and the following warning should accompany the estimates: "Please be warned that these estimates do not meet Statistics Canada's quality standards. Conclusions based on these data will be unreliable, and most likely invalid."

Table 5-2 shows the precision of the SFS estimates. At the Canada level, the estimates are generally reliable. However, users should exercise caution when producing detailed estimates at the regional level.

Additional variance estimates can be calculated by Statistics Canada on a cost-recovery basis.

Table 5-2 Coefficients of variation, Canada and regions, 2005[3]

	All family units					
	million					
	Canada	Atlantic	Quebec	Ontario	Prairies	British Columbia
	%					
Assets (WATOTPT)	**2.75**	**6.98**	**4.87**	**4.50**	**7.40**	**7.01**
RRSPs/LIRAs (WARRSPL)	4.19	11.14	9.24	6.22	13.18	9.20
RRIF(WARRIFF)	10.77	26.66	23.90	17.89	15.09	24.22
EPPs (WARPPT)	3.47	9.01	6.75	5.90	6.09	9.05
Retirement funds, other (WASTOINP)	16.75	31.63	60.82[1]	21.30	16.72	26.46
Deposits in financial institutions (WASTDEPT)	5.53	14.99	12.83	9.27	13.09	11.30
Mutual funds and investment funds (WASTMUIC)	12.35	24.67	22.54	13.85	31.70	15.68
Stocks (WASTSTCK)	13.06	38.74[1]	27.40	17.60	32.50	32.03
Bonds (WASTBOND)	21.85	43.21[1]	36.44[1]	40.77[1]	27.30	48.66[1]
Principal residence (WAPRVAL)	2.00	6.65	4.57	3.01	3.93	5.47
Other real estate (WASTREST)	14.78	20.53	21.35	31.04	26.35	17.63
Vehicles (WASTVHLE)	3.49	7.03	4.84	4.15	4.20	14.92
Other non-financial assets (WASTONOF)	2.97	7.16	5.67	5.04	5.84	10.52
Equity in business (WBUSEQ)	16.11	44.39[1]	31.07	25.68	28.39	38.79[1]
Debts (WDTOTAL)	**3.00**	**5.35**	**6.73**	**4.82**	**5.83**	**7.95**
Mortgage on principal residence (WDPRMOR)	3.49	7.14	8.16	5.85	7.15	6.85
Mortgage on other real estate (WDSTOMOR)	11.81	32.02	26.04	19.15	20.84	32.03
Line of credit (WDSTLOC)	7.14	11.85	19.58	10.74	16.05	16.31
Credit card and instalment debt (WDSTCRED)	7.49	10.65	11.10	15.85	9.16	9.75
Student loans (WDSLOAN)	7.76	23.45	15.93	14.79	13.30	18.78
Vehicle loans (WDSTVHLN)	4.03	9.42	7.55	8.01	7.79	11.10
Other debt (WDSTODBT)	11.73	19.84	18.09	17.75	25.87	32.97
Net worth (WNETWPT)	**3.07**	**8.07**	**5.07**	**5.13**	**8.38**	**7.70**

1. Very high coefficients of variation.

The **bootstrap approach**, a pseudo-replication technique, is used for the calculation of the coefficients of variation of the estimates presented in table 5-2. Many Statistics Canada surveys use complex sampling designs when selecting their samples. As variance estimation for these sampling schemes cannot be accomplished using simple formulae, we must use approximate methods to estimate variances. Resampling methods, and in particular the bootstrap method, figure among these. The bootstrap approach possesses many interesting properties and is the method employed by many Statistics Canada surveys.

3. The coefficients of variation were generated on the internal SFS database, not the public use microdata file.

For more information on the bootstrap approach, refer to the Statistics Canada publication (Catalogue 12-002-XIE), *The Research Data Centres Information and Technical Bulletin*, Fall 2004, vol. 1 no. 2.

5.3 Non-sampling errors

Non-sampling errors can be defined as errors arising during the course of all survey activities other than sampling. Unlike sampling errors, they can be present in both sample surveys and censuses.

Non-sampling errors can be classified into two groups: random errors and systematic errors.

- **Random errors** are the unpredictable errors resulting from estimation. They are generally cancelled out if a large enough sample is used. However, when these errors do take effect, they often lead to an increased variability in the characteristic of interest (i.e., the greater the difference between the population units, the larger the sample size required to achieve a specific level of reliability).

- **Systematic errors** are those errors that tend to accumulate over the entire sample. For example, if there is an error in the questionnaire design, this could cause problems with the respondent's answers, which in turn, can create processing errors, etc. These types of errors often lead to a bias in the final results.

Non-sampling errors are extremely difficult, if not impossible, to measure. Since random errors have the tendency to be cancelled out, systematic errors are the principal cause for concern. Unlike sampling variance, bias caused by systematic errors cannot be reduced by increasing the sample size.

Non-sampling errors can occur because of problems in **coverage**, **response**, **non-response**, **data processing**, **estimation and analysis**.

5.3.1 Coverage errors

An error in coverage occurs when there is an omission, duplication or wrongful inclusion of the units in the population or sample. Omissions are referred to as undercoverage, while duplication and wrongful inclusions are called overcoverage. These errors are caused by defects in the survey frame: inaccuracy, incompleteness, duplication, inadequacy and obsolescence. Coverage errors may also occur in field procedures (e.g., a survey is conducted, but the interviewer misses several households or persons).

5.3.2 Response errors

Response errors result from data that have been requested, provided, received or recorded incorrectly. The response errors may occur because of inefficiencies with the questionnaire, the interviewer, the respondent or the survey process.

Poor questionnaire design
It is essential that sample survey or census questions are worded carefully in order to avoid introducing bias. If questions are misleading or confusing, then the responses may end up being distorted.

Interview bias
An interviewer can influence how a respondent answers the survey questions. This may occur when the interviewer is too friendly or aloof or prompts the respondent. To prevent this, interviewers must be trained to remain neutral throughout the interview. They must also pay close attention to the way they ask each question. If an interviewer changes the way a question is worded, it may impact the respondent's answer.

Respondent errors
Respondents can also provide incorrect answers. Faulty recollections, tendencies to exaggerate or underplay events, and inclinations to give answers that appear more 'socially desirable' are several reasons why a respondent may provide a false answer.

Problems with the survey process
Errors can also occur because of a problem with the actual survey process. Using proxy responses (taking answers from someone other than the respondent) or lacking control over the survey procedures are just a few ways of increasing the possibility for response errors.

5.3.3 Non-response errors

Non-response errors are the result of not having obtained sufficient answers to survey questions. There are two types of non-response errors: complete and partial. The overall response rate for the 2005 Survey of Financial Security was 67.7%.

Complete non-response errors
These errors can occur when the survey fails to measure some of the units in the selected sample. Reasons for this type of error may be that the respondent is unavailable or temporarily absent, the respondent is unable or refuses to participate in the survey, or the dwelling is vacant. If a significant number of people do not respond to a survey, then the results may be biased since the characteristics of the non-respondents may differ from those who have participated.

Partial non-response errors
This type of error deals with incomplete information obtained from the respondent. For certain people, some questions may be difficult to understand. To reduce this form of bias, care should be taken in designing and testing questionnaires. Appropriate edit and imputation strategies will also help minimize this bias.

5.3.4 Processing errors

Processing errors sometimes emerge during the preparation of the final data files. For example, errors can occur while data are being coded, captured, edited or imputed. Coder bias is usually a result of poor training or incomplete instructions, variance in coder performance (i.e., tiredness, illness), data entry errors, or machine malfunction (some processing errors are caused by errors in the computer programs). The same thing can be said about captured errors. Sometimes, errors are incorrectly identified during the editing phase. Even when errors are discovered, they can be corrected improperly because of poor imputation procedures. To minimize errors, diagnostic tests are carried out periodically to ensure that expected results have been obtained.

5.3.5 Estimation errors

Statistics Canada and other data-collecting agencies devote much effort to designing and monitoring surveys in order to make them as error-free as possible. If an inappropriate estimation method is used, then bias can still be introduced, regardless of how errorless the survey had been before estimation.

5.3.6 Analysis error

Analysis errors include any errors that occur when using the wrong analytical tools or when the preliminary results are used instead of the final ones. Errors that occur during the publication of these data results are also considered analysis errors.

5.4 Treatment of large values

For any sample, estimates can be affected disproportionately by the presence or absence of extreme values from the population. In an asset and debt survey, a few extreme values are expected in the sample, as valid

extreme values do exist in the population. Values outside defined bounds were identified and reviewed in relation to other information reported for that respondent. If the value was judged to be the result of a reporting or processing error, it was adjusted. Otherwise, it was retained.

5.5 Impact of sampling and non-sampling errors on SFS estimates

Due to the combined effect of these errors, the quality of net worth data is judged to be lower than the quality of income data. This is largely because records of the current value of assets and the outstanding amount of debt are not as readily available as records of income. For example, respondents with numerous bank accounts and investments may receive several different statements, with different reference periods. Compiling this information can be difficult; most income information, on the other hand, would be available in one document, if the respondent had completed an income tax return for the year in question.

5.6 Comparability of data and related sources

It is important to realize that there are no other sources for much of the data collected by SFS. Of the variables that do have sources, comparison is often difficult because of differences in defining concepts, grouping of items, and how these items are valued.

Direct comparisons with outside sources, such as the National Balance Sheet Accounts (NBSA) of the System of National Accounts (SNA), do yield certain differences. Comparing both of these sources is difficult due to definitional, coverage and treatment differences.

Based on rough comparisons between the NBSA and the SFS, the following general conclusions can be drawn:

(a) The SFS appears to underestimate some net worth components, particularly financial assets and consumer debt.
(b) The quality of estimates of real assets (e.g., owner-occupied homes, vehicles) is much better than that of financial assets.

In theory – given similar valuation procedures and groupings – SNA data should be the same as that collected by an asset and debt survey. The SNA collects individual wealth data from institutional sources such as banks and insurance companies, net of corporations and governments. One major problem has been the SNA categorization of individuals and unincorporated business. Because the individual data and the unincorporated business can not be separated out, these estimates will always be higher than the survey estimates alone.

The Census and other surveys are important sources for ensuring that the SFS sample is representative of the Canadian population. Despite conceptual differences with the SNA estimates, ensuring a representative sample is extremely important to the validity of the data. It was determined that with respect to characteristics such as sex, age, marital status, education that the 2005 SFS data was very comparable to data from the 2001 Census. SFS estimates for pension variables such as membership and contributions were found to be very close to data produced by Statistics Canada's Pension Plans in Canada Survey.

5.7 Response rates

The overall response rate for the 2005 Survey of Financial Security was 67.7%. Table 5-3 gives a breakdown of response rates by province for the area sample and the high income sample.

Table 5-3 Response rates by province

	Area sample	High income sample	Overall
		%	
All provinces	71.2	51.0	67.7
Newfoundland	74.4	48.8	70.4
Prince Edward Island	79.3	58.8	76.6
Nova Scotia	71.6	57.4	69.8
New Brunswick	74.4	70.2	73.8
Quebec	73.2	51.9	69.9
Ontario	64.4	48.2	60.9
Manitoba	74.4	47.6	70.7
Saskatchewan	76.2	46.2	73.6
Alberta	74.2	54.1	69.7
British Columbia	72.6	49.0	69.4

6.0 Record layout, data dictionary and univariate distributions

Three additional information files are provided to assist users of the SFS public use microdata file. A record layout, a data dictionary and univariate distributions are provided. These information files are organized by content themes and in some cases sub-themes.

6.1 Columns of the record layout

Variable name: Public use microdata file (PUMF) variable name: This is the variable name assigned for the microdata file. In almost every case, this name is identical to the name on the SFS internal database.

Type: Indicates whether the variable is numeric (in the sense that it can logically be used in mathematical operations) or character.

Number of categories: Shows the number of categories in the value set for the variable in question. The number applies only to "character" variables. Numeric variables have ranges, which are specified in the data dictionary.

Length: Indicates the number of spaces. For numeric variables, this includes the decimal point if there are decimal places and the number of decimal places, if any. For example, a variable which can have values of zero (00.0) to 99.9 would have a format expressed as 4.1. A variable which can have values of zero (00) to 99 would have a format express as 2.0.

Sequence number: Indicates the order that variables appear on the microdata file.

Start position: This shows the location of the variable on the public use microdata file.

Long variable name: A standardized name, with a maximum of 26 characters, which can be used to quickly identify variables, to label tables, and so on. Although still rather cryptic, it is considerably more revealing than the variable name. However, this longer name obviously excludes a lot of important information contained in the variable description shown in the data dictionary. In short, analysts are warned against making assumptions about the variable definition based on the long variable name.

6.2 Data dictionary

The data dictionary presents the complete information about each survey variable on each of the three files. For each variable are shown: the variable name, the description or definition, code lists with descriptions or alternatively the range of values that the variable can take on, the variable type, its length (or format), and the population to which the variable pertains, i.e. for whom it is applicable.

6.3 Univariate distribution

These distributions are provided to allow users of the public use microdata file to verify totals that they produce. These distributions relate to the public use files and not to the internal database; the distributions will be similar but not identical. To compare the public use file to the internal database, please see Appendices A and B at the end of this user guide.

For character variables, the weighted and unweighted frequencies for each code, including reserved codes, are produced. For numeric variables, the values are broken into several ranges and weighted and unweighted frequencies are provided for each range. The minimum value, the maximum value and the weighted mean (excluding reserved codes) are also provided.

6.4 Reserved codes

It is important to account for reserved codes in any analysis, particularly with numeric variables. If your calculation of means or aggregates seems too high, check to ensure that you have excluded reserved codes from the calculation. With a only a few exceptions, the reserved codes are the highest values permitted according to the length of the variable. A brief explanation of reserved codes is provided below.

7, 97, 9.7, etc. :	Don't know / Not stated (the respondent did not have an answer, or the value was rejected during processing without being replaced)
9, 99, 9.9, etc. :	Not applicable

7.0 Guidelines for analysis and presentation

7.1 Applying weights

The microdata on the public use file are unweighted. It is the responsibility of data users to apply the appropriate weights in any results they wish to produce. If proper weights are not used, the estimates derived from the microdata cannot be considered to be representative of the survey population, and will not correspond to those that would be produced by Statistics Canada. On the SFS PUMF, the weight variable is named WEIGHT.

7.2 Rounding guidelines

Once it has been determined whether the results obtained are reliable, the level of rounding indicates the level of precision that the data can actually support. The following guidelines for rounding should be used:

- Estimates of population sub-groups should be rounded to the nearest hundred units.
- Rates and percentages should be rounded to one decimal point.

Note that all calculations are to be derived from their unrounded components, and then rounded using the normal rounding technique.

In normal rounding, if the first or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, in normal rounding to the nearest 100, the estimate 49,448 would be rounded down to 49,400 and an estimate of 49,252 would be rounded up to 49,300. The figure 1.78% would be rounded to 1.8%.

7.3 Missing values and reserved codes

There are a few types of missing values on the public use file.

If the coverage of a variable does not extend to a certain population sub-group, then there are no valid values for that sub-group and the values that do appear are in the form of 9, 99, 9.9 and so on, which indicates that the variable is not applicable. The coverage of each variable on the file is referred to in the data dictionary as the "population". This also applies to derived variables for which some components have been capped to a certain limit.

For certain records, no valid value is available, although the variable is applicable. Possibly the respondent did not provide the information or it failed an edit in processing, and the value was not imputed. Such missing values appear with a code such as 7, 97, 9.7, and so on depending on the format. For certain variables, the number of missing values has been reduced through imputation. Missing values for the income variables have been entirely imputed, but most other variables may have missing values.

The approach for dealing with missing values of this last kind depends on the type of analysis being carried out and the extent of missing data. Although the end solution may be to exclude the records with missing values from the analysis, a review should first be carried out to assess the impact of missing values on the overall representativeness of the data. Is it possible that a bias results from the missing data — for example, are the (other) characteristics of the people with missing values different from those of the observed part of the sample? It may be necessary to take into account the possible impact in some way. In all cases, analysts should note exclusions of records with missing values in their published results.

8.0 Guidelines for release

Microdata users should apply the rules for assessing data quality, below, to all estimates they produce, and retain only those that satisfy the release criteria. Estimates that do not satisfy the release criteria are not reliable.

8.1 Introduction

The guidelines for release and publication make use of the concept of sampling variability to determine whether estimates obtained from the microdata are reliable. Sampling variability is the error in the estimates caused by the fact that we survey a sample rather than the entire population. The concept of standard error and the related concept of coefficient of variation and confidence interval provide an indication of the magnitude of the sampling variability.

The standard error and coefficient of variation do not measure any systematic biases in the survey data which might affect the estimate. Rather, they are based on the assumption that the sampling errors follow a normal probability distribution. Subject to this assumption, it is possible to estimate the extent to which different samples that have the same design and the same number of observations would give different results. This indicates the margin of error that is likely to be included in the estimates derived from our single sample.

For a detailed description of the measures of sampling variability, see A. Satin and W. Shastry, Survey Sampling: A Non-Mathematical Guide, Statistics Canada, Catalogue 12-602E.

8.2 Minimum sizes of estimates for release

In general, the smaller the sample, the greater the sampling variability. Likewise, estimates of small population subgroups are less reliable than estimates of large population subgroups. The minimum allowable sizes of estimates, also called the release cut-offs, are a quick rule for determining whether an estimate can be released, before applying the more rigorous test that uses the coefficient of variation. The release cut-offs are calculated specifically for the Survey of Financial Security, based on the sample size and the sample design.

The cut-off for the unweighted count must be satisfied:

- Unweighted count: The number of observations must be at least 25. If the unweighted count is less than 25, then the weighted estimate should not be released regardless of the value of its coefficient of variation.

8.3 Hypothesis tests provided by statistical software packages

Microdata users should be aware that the results of hypothesis tests (such as the p values accompanying t statistics or Pearson statistics) that are provided automatically by standard statistical software packages are incorrect for data provided by surveys with a complex survey design, such as Survey of Financial Security. Such packages calculate these test results under the assumption of simple random sampling. That is, they do not take into account the special sample design features of SFS such as stratification, clustering, and unequal selection probabilities. While many of the standard packages can account for the unequal selection probabilities in the production of estimates by allowing the use of weights, these packages do not properly take the sample design into account when producing variance estimates that form part of most test statistics.

To perform hypothesis tests, a two-step method can be employed with the standard statistical software to form the test statistics. First, estimate the characteristics of interest using the weights provided on the

microdata file. Second, obtain approximate variance estimates of these characteristics by rerunning the same software procedure as that used for producing the characteristic estimates but using a scaled weight that consists of the original weight divided by the average of the original weights of all the observations being used in your computations. The quantities calculated in the two steps can then be combined to form test statistics. It must be noted that this method provides only rough approximations to the standard errors.

It should be noted that users of the SFS PUMF cannot readily obtain better design-based variance estimates through the use of statistical software specifically designed for survey data. This is because the design information required by these software packages is not currently available on the SFS data file due to confidentiality considerations. However, better variance estimates can be produced by Statistics Canada on a cost recovery basis.

9.0 Appendix A – SFS production totals VS. PUMF totals

The table below provides a comparison between Statistics Canada's internal SFS database and the public use microdata file.

Table 9-1 Comparison of SFS production totals to SFS PUMF, Canada, 2005

	Production totals				PUMF totals				% Difference (PUMF ÷ Production)-1			
	Total family units	Sum million	Mean	Median	Total family units	Sum million	Mean	Median	Total family units	Sum ($ million)	Mean ($)	Median ($)
			$				$		%			
Assets, termination basis (WATOTPT)	**13,347,678**	**5,622,674**	**421,247**	**229,930**	**13,347,657**	**5,619,809**	**421,033**	**233,000**	**-0.0002**	**-0.0510**	**-0.0508**	**1.3352**
RRSPs/LIRAs (WARRSPL)	6,881,009	466,365	67,776	25,000	6,880,865	466,515	67,799	25,000	-0.0021	0.0322	0.0339	0.0000
RRIF (WARRIF)	1,046,982	104,484	99,796	40,730	1,046,930	104,152	99,484	42,000	-0.0050	-0.3178	-0.3126	3.1181
Employer Pension Plans (WARPPT)	6,490,402	1,038,685	160,034	68,305	6,490,113	1,036,355	159,682	67,500	-0.0045	-0.2243	-0.2200	-1.1785
Retirement funds, other (WAOTPEN)	431,773	22,360	51,786	10,000	431,765	22,379	51,831	9,750	-0.0019	0.0850	0.0869	-2.5000
Deposits in financial institutions (WASTDEPT)	11,612,979	237,325	20,436	3,600	11,612,787	238,411	20,530	3,700	-0.0017	0.4576	0.4600	2.7778
Mutual funds and other (WASTMUIC)	1,640,831	133,753	81,516	24,200	1,640,700	134,366	81,896	29,000	-0.0080	0.4583	0.4662	19.8347
Stocks (WASTSTCK)	1,402,204	120,782	86,137	12,136	1,402,276	120,816	86,157	13,000	0.0051	0.0281	0.0232	7.1193
Bonds (WASTBOND)	1,393,575	34,619	24,842	2,500	1,393,495	34,631	24,852	2,900	-0.0057	0.0347	0.0403	16.0000
Other financial assets, non-pension (WASTOINP)	2,266,821	58,109	25,634	6,000	2,266,711	58,136	25,648	6,000	-0.0049	0.0465	0.0546	0.0000
Principal residence (WAPRVAL)	8,264,953	1,879,657	227,425	180,000	8,264,432	1,879,905	227,469	180,000	-0.0063	0.0132	0.0193	0.0000
Other real estate (WASTREST)	2,142,333	480,828	224,441	85,000	2,142,375	478,901	223,537	87,500	0.0020	-0.4008	-0.4028	2.9412
Vehicles (WASTVHLE)	10,061,590	171,205	17,016	11,557	10,061,386	171,101	17,006	11,500	-0.0020	-0.0607	-0.0588	-0.4932
Other non-financial assets (WASTONOF)	13,347,678	284,675	21,328	10,000	13,347,657	283,947	21,273	10,000	-0.0002	-0.2557	-0.2579	0.0000
Equity in business (WBUSEQ)	2,220,847	589,827	265,586	15,794	2,220,724	590,194	265,767	19,000	-0.0055	0.0622	0.0682	20.2988
Debts (WDTOTAL)	**9,262,940**	**760,188**	**82,068**	**44,500**	**9,262,674**	**759,523**	**81,998**	**45,000**	**-0.0029**	**-0.0875**	**-0.0853**	**1.1236**
Mortgage on principal residence (WDPRMOR)	4,557,145	486,071	106,661	90,000	4,556,832	485,517	106,547	90,000	-0.0069	-0.1140	-0.1069	0.0000
Mortgage on other real estate (WDSTOMOR)	623,990	86,076	137,944	90,000	624,017	86,104	137,983	90,000	0.0043	0.0325	0.0283	0.0000
Line of credit (WDSTLOC)	3,322,734	68,131	20,505	9,000	3,322,465	68,079	20,491	9,250	-0.0081	-0.0763	-0.0683	2.7778
Credit card and instalment debt (WDSTCRED)	5,251,823	25,775	4,908	2,400	5,251,594	25,766	4,906	2,300	-0.0044	-0.0349	-0.0407	-4.1667
Student loans (WDSLOAN)	1,573,892	19,974	12,691	9,000	1,574,257	20,001	12,705	9,000	0.0232	0.1352	0.1103	0.0000
Vehicle loans (WDSTVHLN)	3,449,238	46,105	13,367	11,000	3,449,113	46,112	13,369	11,500	-0.0036	0.0152	0.0150	4.5455
Other debt (WDSTODBT)	1,877,742	28,055	14,941	6,000	1,877,516	27,944	14,883	6,000	-0.0120	-0.3957	-0.3882	0.0000
Net worth, termination basis (WNETWPT)	**13,347,678**	**4,862,486**	**364,295**	**148,350**	**13,347,657**	**4,860,285**	**364,130**	**151,500**	**-0.0002**	**-0.0453**	**-0.0453**	**2.1234**

10.0 Appendix B – SFS totals

The following tables were generated from the SFS production database. Users may use these figures to compare their estimates from the microdata file with Statistics Canada's totals.

Table 10-1 Assets, debts, net worth showing millions of dollars and number of family units, Canada and regions, 2005

	Canada	Atlantic	Quebec	Ontario	Prairies	British Columbia	Canada	Atlantic	Quebec	Ontario	Prairies	British Columbia
	Sum ($ million)						Number of family units (,000)					
Assets (WATOTPT)	5,622,674	277,743	1,066,173	2,353,184	886,982	1,038,593	13,348	986	3,359	4,963	2,218	1,822
Debts (WDTOTAL)	760,188	40,950	125,882	350,971	111,334	131,050	9,263	733	2,315	3,521	1,477	1,217
Net worth, (WNETWPT)[2]	**4,862,486**	**236,793**	**940,290**	**2,002,212**	**775,648**	**907,543**	**13,348**	**986**	**3,359**	**4,957**	**2,218**	**1,822**

Table 10-2 Assets, debts, net worth showing median and average amounts for family units, Canada and regions, 2005

	Canada	Atlantic	Quebec	Ontario	Prairies	British Columbia	Canada	Atlantic	Quebec	Ontario	Prairies	British Columbia
	Median amount for family units holding asset and debt						Average amount for family units holding asset and debt					
Assets, (WATOTPT)	229,930	133,112	154,500	300,890	211,253	316,500	421,247	281,579	317,451	474,169	399,817	570,173
Debts (WDTOTAL)	44,500	30,357	23,500	65,000	45,200	50,000[1]	82,068	55,877	54,374	99,690	75,367	107,674
NET WORTH, (WNETWPT)	**148,350**	**111,445**	**103,701**	**192,161**	**141,819**	**199,746**	**364,295**	**240,063**	**279,970**	**403,448**	**349,632**	**498,228**

1. Estimate has a coefficient of variation of 16.81%